Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 1 February 2008 (MN LATEX style file v2.2)

The Correlation Between Star Formation and 21cm Emission During the Reionization Epoch

J. Stuart B. Wyithe[1], Abraham Loeb[2] & Brian P. Schmidt[3]
[1] *School of Physics, University of Melbourne, Parkville, Victoria, Australia*
[2] *Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA 02138*
[3] *The Research School of Astronomy and Astrophysics, Weston Creek, ACT, Australia*
Email: swyithe@physics.unimelb.edu.au, loeb@cfa.harvard.edu, brian@mso.anu.edu.au

1 February 2008

ABSTRACT

Reionization is thought to be dominated by low mass galaxies, while direct observations of resolved galaxies probe only the most massive, rarest objects. The cross-correlation between fluctuations in the surface brightness of the cumulative Lyα emission (which serves as a proxy for the star formation rate) and the redshifted 21cm signal from neutral hydrogen in the intergalactic medium (IGM), will directly probe the causal link between the production of ionizing photons in galaxies and the reionization of the IGM. We discuss the prospects for detecting this cross-correlation for unresolved galaxies. We find that on angular scales $\lesssim 10'$ detection will be practical using widefield near-IR imaging from space in combination with the forthcoming Mileura Widefield Array - Low Frequency Demonstrator. When redshifted 21cm observations of the neutral IGM are combined with space-based near-IR imaging of Lyα emission, the detection on angular scales $\lesssim 3'$ will be limited by the sensitivity of the 21cm signal, even when a small aperture optical telescope ($\sim$ 2m) and a moderate field of view ($\sim$ 10 square degrees) are used. On scales $\gtrsim 3'$, the measurement of cross-correlation will be limited by the accuracy of the foreground sky subtraction.

Key words: cosmology: diffuse radiation, large scale structure, theory – galaxies: high redshift, intergalactic medium

1 INTRODUCTION

The primary goals for studies of the reionization epoch are to determine the nature of the first generation of galaxies, and to observe the causal link between these galaxies and the ionization state of the intergalactic medium (IGM). At the current time, direct observations of resolved galaxies probe only the most massive, rarest objects (Stark, Loeb & Ellis 2007, and references therein). It has been shown that these massive galaxies should correlate with the redshifted 21cm signal from diffuse neutral hydrogen in the IGM prior to the completion of reionization owing to the biased galaxy formation in over-dense regions (Wyithe & Loeb 2007; Furlanetto & Lidz 2007). However, these massive galaxies are not responsible for the bulk of the ionizing photons that reionized the IGM. Rather, reionization was dominated by low mass galaxies, with luminosities below current detection thresholds (Ellis 2007, and references therein). The emission of these unresolved galaxies should therefore also be correlated with the ionization of the IGM, and by extension, with the redshifted 21cm signal. In this paper we suggest that the cross-correlation between the luminosity density of unresolved Lyα emission and the redshifted 21cm intensity will directly probe the connection between the reionization of the IGM and the star formation rate (and hence the production of ionizing photons). We compute the expected amplitude of this cross-correlation, and discuss the prospects for its detection.

Star formation at high redshift has been studied using fluctuations in unresolved near-IR broad-band emission (e.g. Kashlinsky et al. 2005). Since the fluctuations from star formation at high redshift are superimposed on fluctuations from foreground galaxies at low redshift, these measurements have required subtraction of a model for the fluctuating foreground component. In this paper we discuss removal of the foreground fluctuations statistically using the fact that these are uncorrelated with the redshifted 21cm emission. The measurement of 21cm emission is also subject to a fluctuating foreground, which will be correlated with the foreground in the Lyα observations. However it is proposed as part of upcoming 21cm experiments, that the redshifted 21cm foreground be removed using the smoothness of the spectrum of foreground sources, which will be

compared with the rapid frequency fluctuations of the 21cm signal (Morales et al. 2006). This subtraction method will reveal the narrow-band 21cm fluctuations, but will not allow detection of broad-band 21cm fluctuations. Therefore, rather than consider fluctuations in broad-band flux from high redshift star formation, in this paper we instead discuss narrow-band near-IR observations. The fluctuations in flux within narrow band observations would be dominated by the Lyα line of galaxies in a narrow redshift interval, and would therefore be the appropriate choice for detecting the cross-correlation between the signals.

Any model for the reionization of the IGM must describe the relation between the emission of ionizing photons by stars in galaxies and the ionization state of the intergalactic gas. This relation is non-trivial as it depends on various internal parameters (which may vary with galaxy mass), such as the fraction of the gas within galaxies that is converted into stars and accreting black holes, the spectrum of the ionizing radiation, and the escape fraction of ionizing photons from the surrounding interstellar medium as well as the galactic halo and its immediate infall region (see Loeb 2006 for a review). The relation also depends on intergalactic physics. In regions of the IGM that are overdense, galaxies will be over-abundant because small-scale fluctuations need to be of lower amplitude to form a galaxy when embedded in a larger-scale overdensity (Mo & White 1996). On the other hand, the increase in the recombination rate in over-dense regions counteracts this *galaxy bias*. The process of reionization also contains several layers of feedback. Radiative feedback heats the IGM and results in the suppression of low-mass galaxy formation (Efstathiou, 1992; Thoul & Weinberg 1996; Quinn et al. 1996; Dijkstra et al. 2004). This delays the completion of reionization by lowering the local star formation rate, but the effect is counteracted in over-dense regions by the biased formation of massive galaxies. Most models predict that the sum of these effects is dominated by galaxy bias, and that as a result over-dense regions are reionized first. It follows that the cross-correlation between the star formation rate density and redshifted 21cm emission should be negative, as has been suggested for the cross-correlation between massive galaxies and redshifted 21cm emission (Wyithe & Loeb 2007; Furlanetto & Lidz 2007).

A measurement of the expected anti-correlation between the local star formation rate and the ionization state of the IGM, would provide crucial evidence in favor of the stellar UV reionization model over alternative models in which reionization resulted from decaying particles (Hansen & Haiman 2004; Bierman & Kusenko 2006; Kasuya & Kawasaki 2007; Ripamonti et al. 2007) or from a more diffuse X-ray background (Madau et al. 2004; Ricotti et al. 2005). In this paper we examine the feasibility of making this important measurement based on a simple illustrative model for stellar reionization, described in §2. We then derive the cross-correlation between the star formation rate and 21cm emission in § 3, before discussing the prospects for its detection in § 4. Throughout the paper we adopt the set of cosmological parameters determined by *WMAP* (Spergel et al. 2006) for a flat ΛCDM universe.

2 DENSITY DEPENDENT MODEL OF REIONIZATION

In this paper we compute the relation between the local dark matter overdensity and the brightness temperature of redshifted 21cm emission based on the model described in Wyithe & Loeb (2006). Here we summarize the main features of the model and refer the reader to that paper for more details.

The evolution of the ionization fraction by mass $Q_{\delta,R}$ of a particular region of scale R with overdensity δ (at observed redshift $z_{\rm obs}$) may be written as

$$\begin{aligned}\frac{dQ_{\delta,R}}{dt} &= \frac{N_{\rm ion}}{0.76}\left[Q_{\delta,R}\frac{dF_{\rm col}(\delta,R,z,M_{\rm ion})}{dt}\right.\\ &\left.\quad+(1-Q_{\delta,R})\frac{dF_{\rm col}(\delta,R,z,M_{\rm min})}{dt}\right]\\ &- \alpha_{\rm B}Cn_{\rm H}^0\left(1+\delta\frac{D(z)}{D(z_{\rm obs})}\right)(1+z)^3\,Q_{\delta,R}, \qquad (1)\end{aligned}$$

where $N_{\rm ion}$ is the number of photons entering the IGM per baryon in galaxies, $\alpha_{\rm B}$ is the case-B recombination coefficient, C is the clumping factor (which we assume, for simplicity, to be constant), and $D(z)$ is the growth factor between redshift z and the present time. The production rate of ionizing photons in neutral regions is assumed to be proportional to the collapsed fraction $F_{\rm col}$ of mass in halos above the minimum threshold mass for star formation ($M_{\rm min}$), while in ionized regions the minimum halo mass is limited by the Jeans mass in an ionized IGM ($M_{\rm ion}$). We assume $M_{\rm min}$ to correspond to a virial temperature of 10^4K, representing the hydrogen cooling threshold, and $M_{\rm ion}$ to correspond to a virial temperature of 10^5K, representing the mass below which infall is suppressed from an ionized IGM (Dijkstra et al. 2004). In a region of comoving radius R and mean overdensity $\delta(z)=\delta D(z)/D(z_{\rm obs})$ [specified at redshift z instead of the usual $z=0$], the relevant collapsed fraction is obtained from the extended Press-Schechter (1974) model (Bond et al. 1991) as

$$F_{\rm col}(\delta,R,z)={\rm erfc}\left(\frac{\delta_{\rm c}-\delta(z)}{\sqrt{2\left([\sigma_{\rm gal}]^2-[\sigma(R)]^2\right)}}\right), \qquad (2)$$

where ${\rm erfc}(x)$ is the error function, $\sigma(R)$ is the variance of the density field smoothed on a scale R, and $\sigma_{\rm gal}$ is the variance of the density field smoothed on a scale $R_{\rm gal}$, corresponding to a mass scale of $M_{\rm min}$ or $M_{\rm ion}$ (both evaluated at redshift z rather than at $z=0$). In this expression, the critical linear overdensity for the collapse of a spherical top-hat density perturbation is $\delta_c\approx 1.69$.

Equation (1) may be integrated as a function of δ. At a specified redshift, this yields the filling fraction of ionized regions within the IGM on various scales R as a function of overdensity. We may then also calculate the corresponding 21cm brightness temperature contrast

$$T(\delta,R)=22{\rm mK}(1-Q_{\delta,R})\left(\frac{1+z}{7.5}\right)^{-0.5}\left(1+\frac{4}{3}\delta\right), \qquad (3)$$

where the pre-factor of 4/3 on the overdensity refers to the spherically averaged enhancement of the brightness temperature due to peculiar velocities in over-dense regions (Bharadwaj & Ali 2005; Barkana & Loeb 2005).

3 THE LYα LUMINOSITY DENSITY

The density dependent model described in the previous section may be used to estimate the cross-correlation between star formation rate and the ionization state of the IGM. In this section, we begin by computing the star formation rate. Then, in subsequent sections we estimate the auto-correlation functions for both star formation rate and 21cm brightness temperature, as well as the cross-correlation between star formation rate and 21cm brightness temperature.

The UV-luminosity of galaxies is largest during periods of active star formation. In the dense environments within the high redshift inter-stellar medium the density of neutral hydrogen can be substantial, resulting in absorption of the majority of the UV photons produced. Recombinations in the ionized hydrogen then in turn produce Lyα photons. The Lyα emission from high redshift galaxies is therefore powered by concurrent star formation. In this paper we assume Lyα emissivity to be a proxy for the star formation rate, and so begin by computing the luminosity density of Lyα photons. Given an ionizing photon production rate

$$\log_{10}\left(\frac{\dot{I}}{\rm sec^{-1}}\right) = 53.8 + \log_{10}\left(\frac{\dot{M}}{M_\odot {\rm yr}^{-1}}\right) - 0.0029\left(9+\log_{10}(Z)\right)^{2.5}, \quad (4)$$

where $\dot{M}$ is the star formation rate per comoving Mpc^3, and Z the metalicity of a stellar population with a Salpeter initial mass function, the luminosity of Lyα entering the IGM is

$$\Gamma = 2h_{\rm p}\frac{\nu_\alpha}{3}(1-f_{\rm esc})\mathcal{T}\dot{I}, \quad (5)$$

where $f_{\rm esc}$ is the escape fraction of ionizing photons, $h_{\rm p}$ is Planck's constant and ν_α is the frequency of the Lyα transition. The transmission of Lyα photons through the IGM ($\mathcal{T}$) is less than unity and is discussed below. In the above expressions we evaluate the star formation rate within a region of comoving radius R as

$$\dot{M} = f_{\rm star}\frac{\Omega_{\rm b}}{\Omega_m}\rho_{\rm m}(1+\delta)\left((1-Q_{\delta,R})\frac{dF_{\rm col}(\delta,R,M_{\rm min})}{dt} + Q_{\delta,R}\frac{dF_{\rm col}(\delta,R,M_{\rm ion})}{dt}\right). \quad (6)$$

Here $f_{\rm star}$ is the star formation efficiency, $\Omega_{\rm b}$ and $\Omega_{\rm m}$ are the density parameters in matter and baryons, and $\rho_{\rm m}$ is the average comoving mass-density in the Universe.

It is possible that the mass-function of stars in Lyα emitting galaxies is top-heavy, in which case the Lyα luminosity could be an order of magnitude greater than suggested by equations (4-5). Indeed Dijkstra & Wyithe (2007) have noted that this must be the case due to the the large observed equivalent widths in known Lyα emitters, and the small value of Lyα transmission through the IGM (Dijkstra, Lidz & Wyithe 2007). However Dijkstra & Wyithe (2007) also argue that while top-heavy star formation must be present in many high redshift Lyα emitters, in order to be consistent with additional observations the top-heavy formation phase must last for less than 10% of the star formation time-scale in individual galaxies. As a result, Dijkstra & Wyithe (2007) find that the total Lyα emission is dominated by a normal stellar population when averaged over the full star formation history of galaxies at $z \sim 6$.

3.1 The transmission of Lyα photons through the IGM

Due to the strength of the Lyα resonance, a significant fraction of Lyα flux is absorbed in the infalling IGM surrounding a galaxy (Dijkstra, Lidz & Wyithe 2007). The quantity $\mathcal{T}$ in equation (5) is the transmission of Lyα photons through the IGM, and corresponds to the fraction of Lyα photons leaving the galaxy that propagate to an observer. We assume the absorption of Lyα photons in the IGM to be dominated by neutral hydrogen, with a negligible contribution from dust (due to the low metalicity of the high redshift IGM). We also ignore absorption of Lyα photons by dust within the galaxy due to the low metalicity of high redshift stellar populations. In biased models of reionization, over dense regions are reionized first due to their being regions of greater than average star formation. Thus over-dense regions produce positive fluctuations in the luminosity density of galactic Lyα emission. Conversely, neutral hydrogen is located preferentially in under dense regions, which will therefore be sites of lower Lyα transmission. As a result, the variable transmission of Lyα photons could serve to increase the clustering of Lyα galaxies (McQuinn et al. 2007), and hence to also increase the amplitude of fluctuations in the density of Lyα emission. Recently, Dijkstra, Lidz & Wyithe (2007) have conducted a detailed investigation of the Lyα absorption properties of the IGM surrounding a Lyα emitting galaxy. This work concluded that the ionized IGM introduces significant absorption, and that as a result the Lyα transmission is only weakly dependent on the ionization state of the IGM. In particular, Lyα flux from a galaxy embedded in an HII region rather than in a reionized IGM will be subject to only a small amount of additional absorption due to the damping wing of the Lyα resonance. Rather than introduce a complex model for transmission, in this paper we instead assume the transmission to have the same value for all galaxies, and to be independent of overdensity. As a result we may underestimate the amplitude of Lyα fluctuations. The increased fluctuations introduced by variable transmission would increase the amplitude of the cross-correlation signal between Lyα and redshifted 21cm emission. By assuming constant transmission we therefore arrive at conservative estimates for the detectability of the cross-correlation signal.

3.2 Diffuse Lyα emission from the IGM

In calculating the Lyα luminosity density we have neglected the potential contribution from a re-combining IGM. We now show that this process provides a negligible contribution. To see this we note that at $z = 7$, around 10% of baryons are collapsing inside galaxies per Hubble time, and that some fraction of these form stars ($\sim 30\%$). For every baryon taking part in star formation, around 4000 ionizing photons are produced (e.g. Barkana & Loeb 2001). Most ionizing photons do not escape the galaxy, and each of these produces around 2/3 of a Lyα photon. Of the Lyα photons produced, some ($\sim 70\%$) will be absorbed in the IGM surrounding the galaxy. Hence we find $\sim 4000 \times 0.1 \times 0.3 \times$

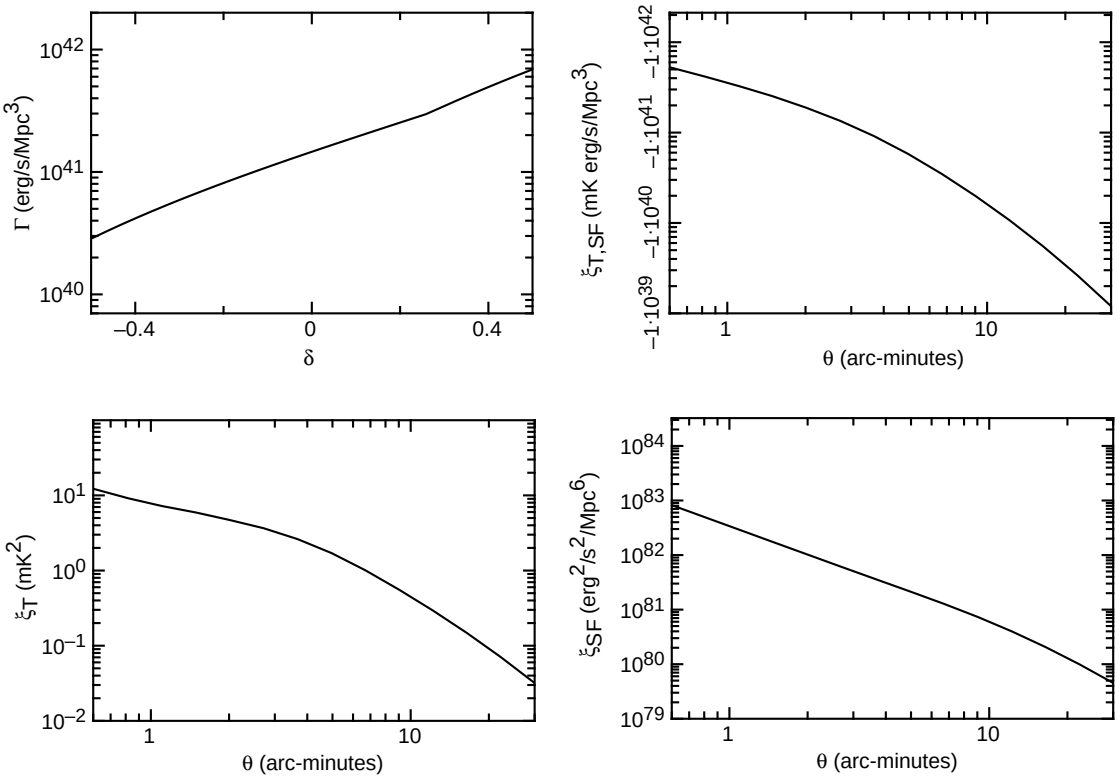

Figure 1. *Upper Left:* The luminosity density (erg per second per comoving Mpc) in the Lyα line as a function of the large scale overdensity (δ). *Upper Right:* The cross-correlation function of the luminosity density in the Lyα line, with the 21cm brightness temperature contrast within spheres of observed radius θ. *Lower Left:* The auto-correlation function of 21cm brightness temperature within spheres of observed radius θ. *Lower Right:* The auto-correlation function of luminosity density in the Lyα line within spheres of observed radius θ.

$(2/3)\times(1-0.7)$, or around 25 photons per baryon are produced by galaxies during 1 Hubble time. On the other-hand, at the redshift of interest, the recombination rate per baryon is around once per Hubble time, yielding of order 2/3 Lyα photons per baryon per Hubble time from the diffuse IGM. This number is 1.5 orders of magnitude smaller than the galactic Lyα emission. A more quantitative estimate of this ratio $R_{\rm Ly}$ is

$$R_{\rm Ly}\sim 50\left(\frac{t_{\rm H}\frac{d\bar{F}_{\rm col}}{dt}}{0.1}\right)\left(\frac{(f_{\rm star}\mathcal{T})N_\gamma}{400}\right)\left(\frac{1+z}{10}\right)^{-\frac{3}{2}}\left(\frac{\bar{Q}}{0.5}\right)^{-1}, \quad (7)$$

where $\bar{Q}$ and $\bar{F}_{\rm col}$ are the average ionized fraction and collapsed fraction in the IGM respectively, and N_γ is the number of ionizing photons produced per baryon incorporated into stars.

4 FIDUCIAL MODEL FOR REIONIZATION

In this paper we show results for the cross-correlation between Lyα and 21cm emission for a model that reionizes the mean IGM at $z=6$ (White et al. 2003). In this model we assume that star formation proceeds in halos above the hydrogen cooling threshold in neutral regions of IGM. In ionized regions of the IGM star formation is assumed to be suppressed by radiative feedback (see § 2). In what follows we present estimates of fluctuations in flux due to sources at $z=7$, at which time the IGM is around 70% ionized in this model. We compute values for auto-correlation functions, and the 21cm-Lyα cross-correlation function at scales as small as $0.6'$. However our model begins to break down on scales below $\sim 1'$, where at $z=7$, 10% of regions have already been reionized on this scale (Wyithe & Morales 2007).

5 VARIATION OF LYα EMISSION WITH OVERDENSITY

The top-left panel of Figure 1 shows the luminosity density (erg per second per comoving Mpc) in the Lyα line as a function of the large scale overdensity (δ). To calculate the level of observed Lyα emission, we require an estimate of the product $f_{\rm star}\mathcal{T}$ (only the product of these parameters enters the observed luminosity). In a recent analysis Dijkstra, Wyithe & Haiman (2007) have used semi-analytic models to constrain this parameter using the observed luminosity function of Lyα emitting galaxies at $z=5.7$ and $z=6.5$. The constraint is sensitive to the life-time of the Lyα emission, but is expected to fall in the range $0.03\lesssim f_{\rm star}\mathcal{T}\lesssim 0.1$. Here, and in the remainder of this paper we assume the product $\mathcal{T}f_{\rm star}=0.1$ when considering the properties of high redshift Lyα emitters.

6 AUTO-CORRELATION FUNCTIONS FOR STAR FORMATION AND 21CM EMISSION

Before discussing the cross-correlation of star formation rate (Lyα emission) with 21cm emission, we first compute each of the auto-correlation functions individually. On comoving scales R larger than the characteristic bubble size ($\gtrsim 1'$ at $z = 7$ in our model), we are able to compute the auto-correlation function $[\xi_T(\theta)]$ of fluctuations in brightness temperature T smoothed with top-hat windows of angular radius $\theta = R/D_{\rm A}(z)$,

$$\begin{aligned}\xi_T(\theta) &= \langle (T - \langle T\rangle)^2\rangle^{1/2} \\ &= \left[\frac{1}{\sqrt{2\pi}\sigma(R)}\int d\delta\,(T(\delta) - \langle T\rangle)^2\, e^{-\frac{\delta^2}{2\sigma(R)^2}}\right]^{\frac{1}{2}}. \quad (8)\end{aligned}$$

Here

$$\langle T\rangle = \frac{1}{\sqrt{2\pi}\sigma(R)}\int d\delta\; T(\delta) e^{-\frac{\delta^2}{2\sigma(R)^2}}, \quad (9)$$

and $\theta = R/D_{\rm A}$ where $D_{\rm A}$ is the angular diameter distance. The auto-correlation function of 21cm brightness temperature within spheres of observed radius θ is plotted in the lower-left panel of Figure 1.

We also compute the auto-correlation function $[\xi_{\rm SF}(\theta)]$ of fluctuations in Lyα emission Γ smoothed with top-hat windows of angular radius $\theta = R/D_{\rm A}(z)$,

$$\begin{aligned}\xi_{\rm SF}(\theta) &= \left\langle \left(\Gamma - \langle\Gamma\rangle\right)^2\right\rangle^{1/2} \\ &= \left[\frac{1}{\sqrt{2\pi}\sigma(R)}\int d\delta\,(\Gamma - \langle\Gamma\rangle)^2\, e^{-\frac{\delta^2}{2\sigma(R)^2}}\right]^{\frac{1}{2}}, \quad (10)\end{aligned}$$

where

$$\langle\Gamma\rangle = \frac{1}{\sqrt{2\pi}\sigma(R)}\int d\delta\; \Gamma(\delta) e^{-\frac{\delta^2}{2\sigma(R)^2}}. \quad (11)$$

The auto-correlation function of luminosity density in the Lyα line within spheres of observed radius θ is shown in the lower-right panel of Figure 1.

7 THE CROSS-CORRELATION FUNCTION BETWEEN LYα AND 21CM EMISSION

The properties of the galaxy population are expected to correlate with the level of redshifted 21cm emission. These properties depend on the overdensity of the IGM whose typical fluctuation level is a function of scale. As a result, the amplitude of the correlation between fluctuations in Lyα emission $(\Gamma - \langle\Gamma\rangle)$ and fluctuations in 21cm brightness temperature contrast $(T - \langle T\rangle)$ will therefore also be dependent on angular scale. On a comoving scale R larger than the characteristic bubble radius, we are able to compute the cross-correlation function between Lyα and 21cm emission

$$\begin{aligned}\xi_{\rm T,SF}(\theta) &= \langle(\Gamma - \langle\Gamma\rangle)\times(T - \langle T\rangle)\rangle \\ &= \frac{1}{\sqrt{2\pi}\sigma(R)}\int d\delta\,((\Gamma - \langle\Gamma\rangle)\times(T - \langle T\rangle))\, e^{\frac{-\delta^2}{2\sigma(R)^2}} \quad (12)\end{aligned}$$

for the IGM smoothed on various angular scales. The resulting cross-correlation function of the luminosity density in the Lyα line, with the 21cm brightness temperature contrast within spheres of observed radius θ is shown in the top-right panel of Figure 1. The sign of this cross-correlation is negative, indicating an anti-correlation between star formation and 21cm emission. This anti-correlation arises as a result of the higher star formation rates generated due to galaxy bias in overdense regions, which are therefore reionized first. The amplitude of the cross-correlation decreases towards large scales.

8 DETECTABILITY OF THE CROSS-CORRELATION SIGNAL

In the remainder of this paper we discuss detection of the predicted cross-correlation between Lyα and the 21cm emission. We begin with the Lyα signal and extra-galactic foreground, which we assume are measured in a wide-field near-IR survey through a narrow-band filter centered on the redshifted Lyα wavelength. We then discuss the sensitivity of planned low-frequency arrays to the redshifted 21cm signal, before describing prospects for detection of the predicted cross-correlation between Lyα and redshifted 21cm emission using a range of current and future observational facilities.

8.1 The Lyα flux

Equation (8) can be used to compute the fluctuations in Lyα luminosity from spherical regions subtending an angle θ,

$$\Delta L_{\rm Ly} = (\xi_{\rm SF})^{1/2}\,\frac{4\pi(\theta D_{\rm A})^3}{3}, \quad (13)$$

while the corresponding total luminosity follows from equation (5)

$$L_{\rm Ly} = \langle\Gamma\rangle\frac{4\pi(\theta D_{\rm A})^3}{3}. \quad (14)$$

For a telescope of diameter d, the fluctuations in the observed photon count are

$$\Delta\dot{N}_{\rm Ly} = \pi\left(\frac{d}{2}\right)^2\frac{\Delta L_{\rm Ly}}{4\pi D_{\rm L}^2}\frac{1}{h_{\rm p}\nu_{\rm obs}}, \quad (15)$$

where $D_{\rm L}$ is the luminosity distance and $\nu_{\rm obs}$ is the observed frequency of the Lyα photons. Similarly, the total flux in Lyα photons is

$$\dot{N}_{\rm Ly} = \pi\left(\frac{d}{2}\right)^2\frac{L_{\rm Ly}}{4\pi D_{\rm L}^2}\frac{1}{h_{\rm p}\nu_{\rm obs}}. \quad (16)$$

Figure 2 shows the observed fluxes and fluctuations in observed fluxes in units of photons per hour. In Figure 2 we assumed a 2m telescope with a 1 hour integration time. The large black dots and thin solid line refer respectively to the fluctuation level ($\Delta\dot{N}_{\rm Ly}$) and the total level ($\dot{N}_{\rm Ly}$) of Lyα emission within a spherical region of observed radius θ.

8.2 Foreground emission

Observations through a narrow filter will detect fluctuations in Lyα emission superimposed on fluctuations in the extra-galactic foreground. To estimate the importance of the foreground with respect to measurement of the cross-correlation we therefore need to estimate both the total foreground flux (F), and the fluctuations in total flux that enters the detector from a cone of angular radius θ at a frequency $\nu_{\rm obs}$. For

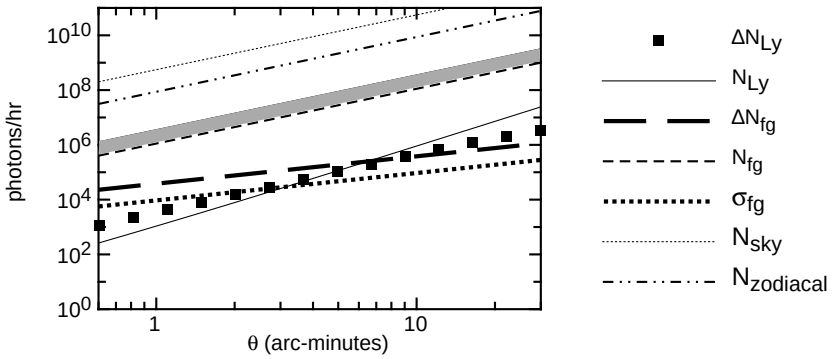

Figure 2. The observed fluxes and fluctuations in observed fluxes in units of photons per hour. Calculation of the flux was performed assuming a 2m telescope and a 1 hour integration. The large black dots and thin solid line refer respectively to the fluctuation level ($\Delta\dot{N}_{\rm Ly}$) and the total level ($\dot{N}_{\rm Ly}$) of Lyα emission within a spherical region of observed radius θ. The thin dashed curve corresponds to the flux ($\dot{N}_{\rm fg}$) in a 100Å band due to foregrounds at the wavelength of the observed Lyα emission. For comparison we also plot the measured extra-galactic foreground at 8000Å as the grey band. The level of fluctuations ($\Delta\dot{N}_{\rm fg}$) in $\dot{N}_{\rm fg}$ among different lines of sight due to Poisson noise in the number of galaxies contributing to the background is shown as the thick dashed line. The levels of sky-glow ($\dot{N}_{\rm sky}$) and zodiacal light ($\dot{N}_{\rm zodiacal}$) are shown by the thin dotted and dot-dot-dashed lines respectively. Finally the thick dotted line shows the Poisson noise ($\sigma_{\rm fg}=\sqrt{\dot{N}_{\rm fg}+\dot{N}_{\rm Ly}+\dot{N}_{\rm zodiacal}}$) in the number of photons detected in a space-based observation (i.e. no sky-glow) per region of radius θ.

our purposes it is sufficient to estimate the foreground flux using the following simple model

$$F(\theta,\nu_{\rm obs}) = \int_0^{z_{\rm Ly}} dz\pi[D_{\rm A}(z)\theta]^2\frac{cdt}{dz}\left.\frac{d^2E}{dVd\nu}\right|_{\nu=\nu_{\rm obs}(1+z)}\frac{(1+z)}{4\pi D_{\rm L}^2}, \quad (17)$$

where

$$\frac{d^2E}{dVd\nu}=f_\star\frac{\Omega_{\rm b}}{\Omega_{\rm m}}\rho_{\rm m}(1+z)^3\frac{dF_{\rm col}}{dt}\frac{d^2E}{d\dot{M}d\nu} \quad (18)$$

is the luminosity density. In the latter expression, $\frac{d^2E}{d\dot{M}d\nu}$ is the luminosity produced at frequency ν per unit star formation rate. We assume a 1/20th solar metalicity population with a Scalo (1998) mass-function, and use the stellar population model of Leitherer et al. (1999) to compute the spectrum of a continuously star forming galaxy[1]. For a narrow filter of width $\Delta\nu_{\rm Ly}$, the flux can then be converted into a photon detection rate

$$\dot{N}_{\rm fg}=\pi\left(\frac{d}{2}\right)^2\frac{F(\theta,\nu_{\rm obs})}{h_{\rm p}\nu_{\rm obs}}\Delta\nu_{\rm Ly}, \quad (19)$$

where $h_{\rm p}$ is Planck's constant. The thin dashed curve in Figure 2 corresponds to the flux in a 100Å band due to foregrounds at the wavelength of the observed Lyα emission from galaxies at $z=7$. This model can be compared to the measured extra-galactic foreground at 8000Å. Bernstein, Freedman & Madore (2002) find flux at a level of (1.2-3)$\times10^{-9}$erg/s/cm^2/Å/Sr. In the units of Figure 2, this observation is shown as the grey band. Our model estimate lies on the lower boundary of the measured range for the observed foreground.

There will be fluctuations ($\Delta\dot{N}_{\rm fg}$) in $\dot{N}_{\rm fg}$ among different lines of sight due to Poisson noise in the number of galaxies contributing to the foreground. The level of fluctuations is given by

$$\frac{\Delta\dot{N}_{\rm fg}}{\dot{N}_{\rm fg}} = \frac{\sqrt{\int_0^{z_{\rm Ly}}dz\int_{M_{\rm ion}}^{M_{\rm lim}(z)}dM\epsilon_{\rm lt}\frac{dn}{dM}\pi\left[D_{\rm D}(z)\theta\right]^2\frac{cdt}{dz}\left[\dot{\mathcal{N}}(M,z)\right]^2}}{\int_0^{z_{\rm Ly}}dz\int_{M_{\rm ion}}^{M_{\rm lim}(z)}dM\epsilon_{\rm lt}\frac{dn}{dM}\pi\left[D_{\rm D}(z)\theta\right]^2\frac{cdt}{dz}\dot{\mathcal{N}}(M,z)} \quad (20)$$

where $\dot{\mathcal{N}}(M,z)$ is the observed flux from a galaxy of mass M at redshift z, and $\epsilon_{\rm lt}$ is the duty-cycle. The presence of bright, resolved galaxies at low redshift increase the fluctuations in the smoothed foreground. To reduce the amplitude of fluctuations in the foreground, these resolved galaxies need to be removed. To estimate the foreground fluctuations in the absence of resolved galaxies, we therefore compute the flux corresponding to a photon limited signal-to-noise ratio of 30, given an assumed telescope diameter, integration time and filter width. As a function of redshift, we then estimate the star formation rate ($\dot{M}_{\rm lim}$) corresponding to this limiting flux. By assuming a star formation efficiency and lifetime (we take $f_{\rm star}=0.1$ and $\epsilon_{\rm lt}=1$ respectively, which is appropriate for the old stellar populations contributing to the foreground), we estimate the limiting galaxy mass, $M_{\rm lim}=\dot{M}_{\rm lim}t_{\rm lt}f_{\rm star}^{-1}\Omega_{\rm m}/\Omega_{\rm b}$. Only galaxies whose masses are below this limit contribute to the unresolved foreground. The upper limits on the mass integration in equation (20) therefore correspond to the minimum galaxy mass that can be removed from the map as a resolved source prior to the calculation of fluctuations.

Figure 2 shows that the level of foreground flux (thin dashed line) is larger than the flux due to galactic Lyα emission at $z\sim7$ (thin solid line) by several orders of magnitude. However the absolute levels of fluctuation are similar (thick dashed and large dotted lines), particularly at larger angular

[1] Model spectra of star forming galaxies obtained from http://www.stsci.edu/science/starburst99/.

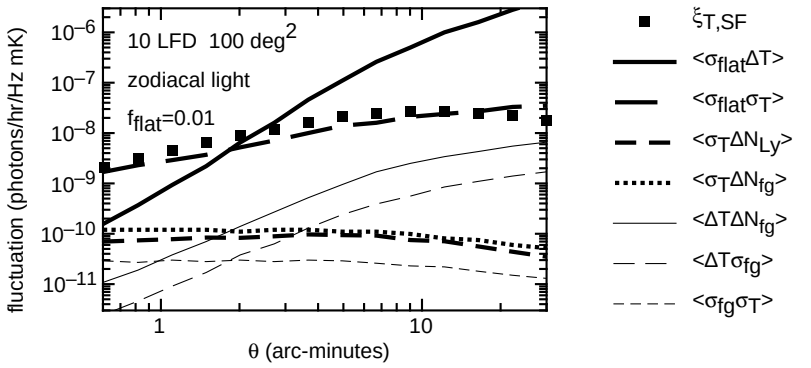

Figure 3. The signal and noise terms in equation (23) as a function of θ. This example assumed a Lyα survey with an area of $A = 100$ square degrees using a 2 meter space-based telescope and 1hr integration per pointing, combined with a low-frequency array of collecting area 10 times the LFD with an integration time of 1000 hours. The near-IR observations were assumed to be flattened at the 1% level.

scales[2]. The fluctuations in the foreground and Lyα emission have different angular dependencies because the foreground fluctuations are dominated by Poisson noise, while the Lyα fluctuations are a result of the biased star formation rates toward over-dense regions.

In addition to the extra-galactic foreground generated by stellar continuum, we expect a fluctuating foreground due to emission lines at frequencies blueward of Lyα, from sources located at redshifts below the Lyα emitting galaxies. Like the Lyα fluctuations, these will have large relative amplitudes due to the narrow redshift bin in which the sources contribute to the foreground. However unlike the Lyα fluctuations this foreground of line emission will not correlate with the redshifted 21cm emission. Since, as we show below, the fluctuating foreground does not provide the limiting factor in detection of the correlation between Lyα emission and 21cm intensity, we therefore neglect the contribution of low redshift emission line sources to the extra-galactic foreground in the remainder of this paper.

However images will be contaminated with additional foregrounds due to zodiacal light, and, for ground based observation, due to atmospheric sky glow in addition to the extragalactic foreground. In Figure 2 we show the level of sky-glow based on the Keck skyglow spectrum ($\dot{N}_{\rm sky}$) as the thin dotted line, and the level of zodiacal light measured at 9000Å ($\dot{N}_{\rm zodiacal}$) towards the ecliptic pole (Leinert et al. 1997) as the dot-dot-dashed line. Zodiacal light is around 2 orders of magnitude larger, and atmospheric skyglow around 3 orders of magnitude larger than the extragalactic foreground.

In order to detect the fluctuations in the correlation between Lyα and 21cm emission, the observational noise in the foreground must be smaller than the fluctuations being measured. The thick dotted line in Figure 2 shows the Poisson noise ($\sigma_{\rm fg} = \sqrt{\dot{N}_{\rm fg} + \dot{N}_{\rm Ly} + \dot{N}_{\rm zodiacal}}$) in the number of photons detected per region of radius θ (including zodiacal light but not atmospheric sky-glow). For the example shown the Poisson noise is comparable to the size of fluctuations in Lyα emission.

[2] Note that the flux levels of the foreground and the Lyα signal shown in Figure 2 have different power-law dependences on θ. This is because the foreground has been calculated in a cone, while the Lyα emission has been calculated in spheres.

8.3 Near-IR field flatness

Finally, we mention one further source of fluctuations that could mask the fluctuations in Lyα emission. Variability of the foregrounds in time and space can be removed through dithering techniques to produce a foreground free, and flattened field containing only the differential fluctuations in the signal (plus extragalactic foregrounds). However the field can only be flattened to a finite fractional level (e.g. 0.01), which we define to be $f_{\rm flat}$. Observations will therefore contain an additional fluctuating term with an amplitude of $\sigma_{\rm flat} = f_{\rm flat}(\dot{N}+\dot{N}_{\rm fg}+\dot{N}_{\rm sky}+\dot{N}_{\rm zodiacal})$. We will find that this term dominates the error budget on angular scales greater than a few arc-minutes, and in the case of ground based observations, that it will prevent detection of the cross-correlation on those scales.

8.4 Sensitivity to the 21cm signal

In this section we discuss the response of a phased array to the brightness temperature contrast of the 21cm emission from the IGM. We define the error in brightness temperature per synthesized beam to be $\sigma_{\rm T}$. Assuming that calibration can be performed ideally, and that redshifted 21cm foreground subtraction is perfect, the root-mean-square fluctuations in brightness temperature are given by the radiometer equation

$$\sigma_{\rm T} = \frac{\epsilon\lambda^2 T_{\rm sys}}{A_{\rm tot}\Omega_{\rm b}\sqrt{t_{\rm int}\Delta\nu_{21}}}, \qquad (21)$$

where λ is the wavelength, $T_{\rm sys}$ is the system temperature, $A_{\rm tot}$ the collecting area, $\Omega_{\rm b}$ the effective solid angle of the synthesized beam in radians, $t_{\rm int}$ is the integration time, $\Delta\nu_{21}$ is the size of the frequency bin, and ϵ is a constant that describes the overall efficiency of the telescope. We optimistically adopt $\epsilon = 1$ in this paper. In units relevant for upcoming telescopes and at $\nu = 200$MHz, we find (Wyithe,

Loeb & Barnes 2005)

$$\sigma_{\rm T} = 7.5{\rm mK}\left(\frac{1.97}{C_{\rm beam}}\right)\left(\frac{A_{\rm tot}}{A_{\rm LFD}}\right)^{-1} \times \left(\frac{\Delta\nu_{21}}{1{\rm MHz}}\right)^{-1/2}\left(\frac{t_{\rm int}}{100{\rm hr}}\right)^{-1/2}\left(\frac{\theta_{\rm beam}}{5'}\right)^{-2}. \quad (22)$$

The label *LFD* corresponds to the Low-Frequency Demonstrator of the Mileura Wide-Field Array (see http://www.haystack.mit.edu/ast/arrays/mwa/site/index.html). $A_{\rm LFD}$ is the collecting area of a phased array consisting of 500 tiles each with 16 cross-dipoles [the effective collecting area of an LFD tile with 4×4 cross-dipole array with 1.07m spacing is $\sim$ 17–19m^2 between 100 and 200MHz (B. Correy, private communication)]. The system temperature at 200MHz will be dominated by the sky and has a value $T_{\rm sys} \sim 250$K. The size of the synthesized beam $\theta_{\rm beam}$ can be regarded as the radius of a hypothetical top-hat beam, or as the variance of a hypothetical Gaussian beam. The corresponding values of the constant $C_{\rm beam}$ are 1 and 1.97 respectively.

8.5 Estimate of signal-to noise ratio in detection of the cross-correlation

The observed cross-correlation function ($\xi_{\rm Ly,T}$) is a combination of real fluctuations and noise, hence we can write

$$\begin{aligned}\xi^{\rm obs}_{\rm Ly,T} &= \langle\left(\Delta\dot{N}_{\rm Ly}+\Delta\dot{N}+\sigma_{\rm fg}+\sigma_{\rm flat}\right)(\Delta T+\sigma_{\rm T})\rangle \\ &= \xi_{\rm Ly,T}+\langle\Delta\dot{N}_{\rm Ly}\sigma_{\rm T}\rangle+\langle\Delta\dot{N}\Delta T\rangle+\langle\Delta\dot{N}\sigma_{\rm T}\rangle \\ &\quad +\langle\sigma_{\rm fg}\Delta T\rangle+\langle\sigma_{\rm fg}\sigma_{\rm T}\rangle+\langle\sigma_{\rm flat}\Delta T\rangle+\langle\sigma_{\rm flat}\sigma_{\rm T}\rangle. \quad (23)\end{aligned}$$

Here we have assumed prior removal of spectrally smooth foreground from the redshifted 21cm maps (this removal is expected to be part of the real time data processing pipeline for an instrument like the LFD). We have also defined $\Delta T \equiv T - \langle T\rangle$. The fluctuations and noise in the foreground should be un-correlated with the 21cm signal. Similarly, the noise in the 21cm signal should be uncorrelated with each of the Lyα fluctuations, the noise in Lyα flux, and the level of foreground. Terms 2-8 in the above equation therefore average individually to zero over a large sample. However for a finite sample, the expectation value will have a distribution with a finite variance about zero. To examine the variance, consider two variables x and y. Their product has a distribution $p(xy)$ with variance σ_{xy}. If we sample this distribution $N_{\rm points}$ times, the resulting mean is distributed about zero with a variance $\langle xy\rangle = \sigma_{xy}/\sqrt{N_{\rm points}}$. Since 21cm surveys are inherently wide-field, the number of independent terms in the cross-correlation will be limited by optical surveys. The width of the frequency bin $\Delta\nu_{21}$ corresponds to the line-of-sight depth of a spherical region of radius θ. At small angles this depth can be smaller than the line-of-sight distance corresponding to a narrow (100Å) near-IR band. Thus if a map of Lyα emission has an area $A_{\rm sky}$, then the number of regions is

$$N_{\rm points} \sim \frac{A_{\rm sky}}{\pi\theta^2}\left(\frac{\Delta\nu_{\rm Ly}/\nu_{\rm Ly}}{\Delta\nu_{21}/\nu_{21}}\right), \quad (24)$$

where ν_{21} and $\nu_{\rm Ly}$ are the redshifted frequencies of the 21cm and Lyα emission respectively.

Figure 3 shows each of the noise terms in equation (23) as a function of θ, corresponding to the case of a Lyα survey with an area of $A_{\rm sky} = 10$ square degrees flattened at the 1% level ($f_{\rm flat} = 0.01$), with an LFD integration time of 1000 hours. Also shown is the expected cross-correlation function (large dots). In the case shown, the cross correlation would be only marginally detectable. The figure demonstrates that at large angular scales the detection is limited by the flatness of the Lyα field achieved in the experiment. At small angular scales the detection is limited by the error in the brightness temperature of a synthesized 21cm beam. Improved measurements would therefore require flatter fields and larger radio arrays rather than deeper near IR imaging.

The signal-to-noise ratio for detection of the cross-correlation is given by

$$(SN)^2 = \frac{(\xi_{\rm Ly,T})^2}{\Sigma^2} \quad (25)$$

where

$$\begin{aligned}\Sigma^2 &= \langle\Delta\dot{N}_{\rm Ly}\sigma_{\rm T}\rangle^2+\langle\Delta\dot{N}\Delta T\rangle^2+\langle\Delta\dot{N}\sigma_{\rm T}\rangle^2 \\ &\quad +\langle\Delta T\sigma_{\rm fg}\rangle^2+\langle\sigma_{\rm flat}\sigma_{\rm T}\rangle^2+\langle\sigma_{\rm flat}\sigma_{\rm T}\rangle^2. \quad (26)\end{aligned}$$

Signal-to-noise ratios as a function of angle are plotted in Figure 4 assuming parameters corresponding to a range of observational facilities. Six cases are shown in each panel, corresponding to Lyα surveys with areas of $A_{\rm sky} =$ 10 and 100 square degrees performed using a 2m telescope with 1 hour of integration per pointing; combined with low-frequency arrays of collecting area corresponding to 1, 10 and 100 LFDs with an integration time 1000 hours. The latter example of a low-frequency array has around a square kilometer of collecting area and would represent the realisation of a square kilometer array (see http://www.skatelescope.org/). The left and right panels correspond to space based (i.e. no atmospheric skyglow, but including zodiacal light), and ground based (i.e. including skyglow) imaging. The 2m space based telescope capable of widefield imaging might represent a telescope like the planned Supernova Acceleration Probe (SNAP[3]). The value of $f_{\rm flat}$ (ranging between 0 and 0.01) is listed in each case. We have shown examples with values of $f_{\rm flat}$ that are an order of magnitude smaller for ground based examples. The upper row with $f_{\rm flat} = 0$ represents an experiment with a perfectly flat near-IR image field.

Figure 4 shows that the cross correlation will be detectable at angles below a few arc-minutes using wide-field space based imaging ($\sim$ 100 square degrees) combined with 10 times the LFD, provided that the Lyα images can be flattened at the $\sim$ 0.1% level. Ground based studies will be limited by the flatness of the near-IR imaging field, and would need to reach values of $f_{\rm flat} \sim 10^{-4}$ over 100 square degrees. At large angles the signal to noise ratio is limited by the value of $f_{\rm flat}$, and by the area of the survey. However at small angles the measurement of cross-correlation is limited by the noise in the 21cm observations, and so the signal-to-noise ratio is proportional to collecting area of the low-frequency array. The greater sensitivity of an SKA therefore increases the signal to noise of the detection on scales near an arc-minute. For example, a signal-to-noise ratio greater

[3] see http://snap.lbl.gov/

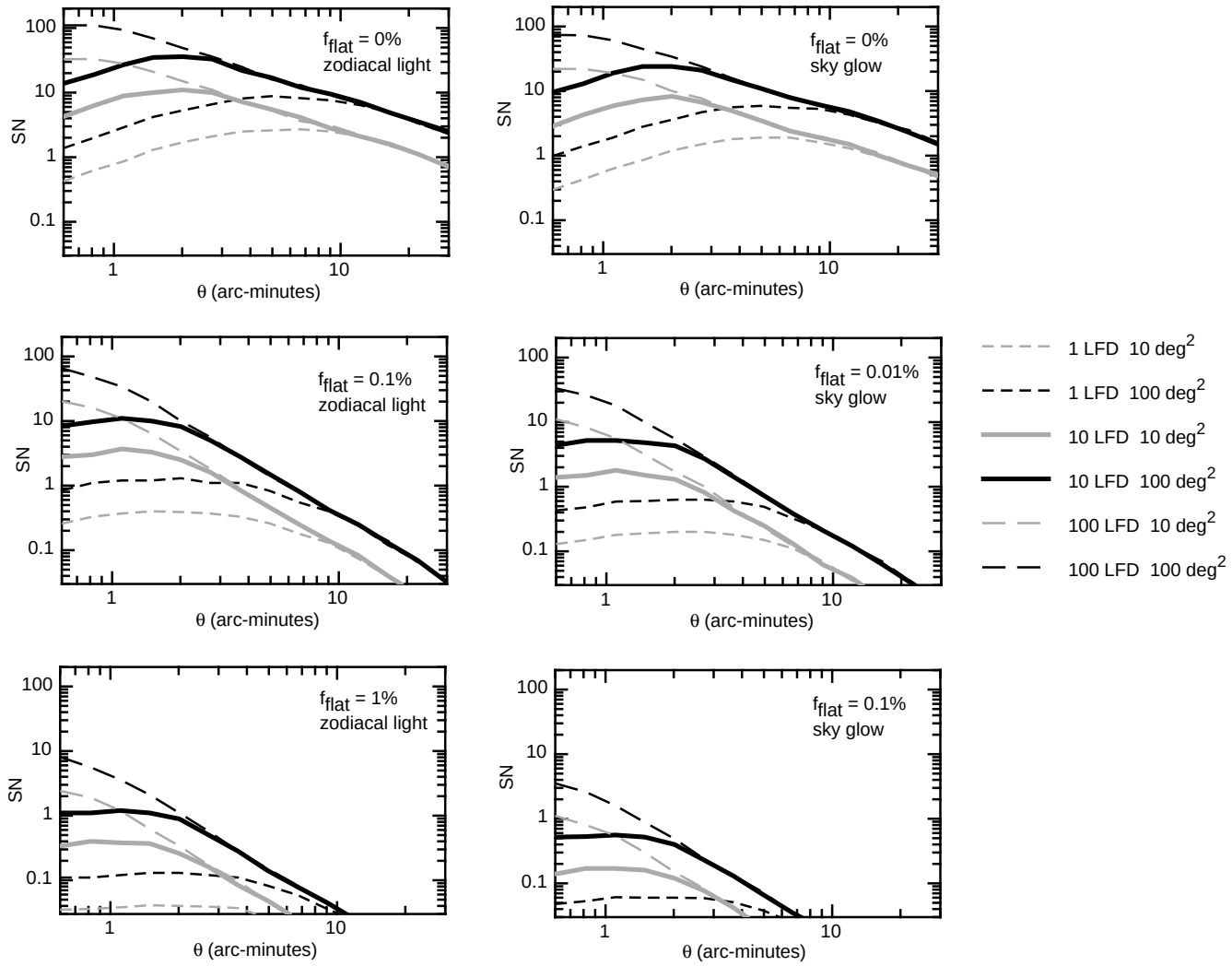

Figure 4. Signal to noise ratios as a function of angle. In each panel six cases are shown, corresponding to Lyα surveys with areas of $A = 10$ and 100 square degrees using a 2m telescope and a 1 hour integration; combined with low-frequency arrays of collecting area corresponding to 1, 10 and 100 LFDs with an integration time 1000 hours. The left-hand panels correspond to space based (i.e. no sky-glow, but including zodiacal light), and the right-hand panels to ground based near-IR observations (i.e. including sky glow). The value of $f_{\rm flat}$ is listed in each case. Note the assumed values for $f_{\rm flat}$ are an order of magnitude lower for ground based observations.

than 10 could be achieved in a 100 square degree survey combined with a ground based near-IR survey with $f_{\rm flat} = 10^{-4}$ or with a space based near-IR survey with $f_{\rm flat} = 0.01$.

8.6 Signal-to noise ratio in the case of uniform zodiacal light

The signal-to-noise ratio results presented thus far have assumed that the sky-glow and zodiacal light leave an imprint on the measured fluctuations via the instrumental effect of an imperfect flat-field. However zodiacal light has very small spatial and temporal fluctuations (Kashlinsky, Arendt, Mather & Moseley 2007). In space based observations, the fluctuations introduced by variable instrumental response to the zodiacal light could therefore be removed via subtraction of two independent regions of sky, which we label A and B. Before concluding this paper, we therefore estimate the signal-to-noise ratio of an analysis conducted this way.

As in equation (23), the observed cross-correlation function ($\xi_{\rm Ly,T}$) is a combination of real fluctuations and noise, hence we can write the cross-correlation of differences between two regions of the map, *measured using the same area of detector* as

$$\begin{aligned}
&\frac{1}{2}\langle\Big[\Big(\Delta\dot{N}^{\rm A}_{\rm Ly}+\Delta\dot{N}^{\rm A}+\sigma^{\rm A}_{\rm fg}+\sigma_{\rm flat}\Big)\\
&\qquad-\Big(\Delta\dot{N}^{\rm B}_{\rm Ly}+\Delta\dot{N}^{\rm B}+\sigma^{\rm B}_{\rm fg}+\sigma_{\rm flat}\Big)\Big]\\
&\qquad\times\Big[\Big(\Delta T^{\rm A}+\sigma^{\rm A}_{\rm T}\Big)-\Big(\Delta T^{\rm B}+\sigma^{\rm B}_{\rm T}\Big)\Big]\rangle\\
=\;&\xi_{\rm Ly,T}+\langle\Delta\dot{N}^{\rm A}_{\rm Ly}\Delta T^{\rm B}\rangle\\
&\qquad+2\Big[\langle\Delta\dot{N}_{\rm Ly}\sigma_{\rm T}\rangle+\langle\Delta\dot{N}\Delta T\rangle+\langle\Delta\dot{N}\sigma_{\rm T}\rangle\\
&\qquad\qquad+\langle\sigma_{\rm fg}\Delta T\rangle+\langle\sigma_{\rm fg}\sigma_{\rm T}\rangle\Big] \qquad (27)
\end{aligned}$$

In obtaining the second line of equation (27), we have noted

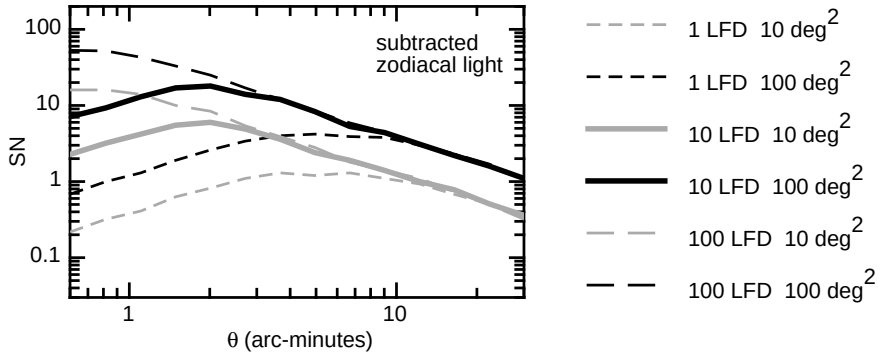

Figure 5. Signal to noise ratios as a function of angle. Six cases are shown, corresponding to Lyα surveys with areas of $A = 10$ and 100 square degrees using a 2m telescope and a 1 hour integration; combined with low-frequency arrays of collecting area corresponding to 1, 10 and 100 LFDs with an integration time 1000 hours. The simulation assumes space based observation (i.e. no sky-glow, but including zodiacal light), where the contribution to fluctuations due to imperfections in the flatness of the field are removed via subtraction of the constant fore-ground which is dominated by zodiacal light.

that if the zodiacal light is constant then the contribution from terms containing $\sigma_{\rm flat}$ disappears[4]. We have also used the fact that the product of un-correlated quantities between different regions A and B has the same distribution as the product obtained from the same region (A or B). In this case the signal-to-noise ratio for detection of the correlation is given by

$$(SN)^2 = \frac{(\xi_{\rm Ly,T})^2}{\Sigma^2} \quad (28)$$

where

$$\Sigma^2 = \langle \Delta\dot{N}^{\rm A}_{\rm Ly}\Delta^{\rm B}_{\rm T}\rangle^2 + 4\Big[\langle\Delta\dot{N}_{\rm Ly}\sigma_{\rm T}\rangle^2 + \langle\Delta\dot{N}\Delta T\rangle^2 + \langle\Delta\dot{N}\sigma_{\rm T}\rangle^2 + \langle\Delta T\sigma_{\rm fg}\rangle^2\Big] \quad (29)$$

Signal to noise ratios as a function of angle are plotted in Figure 5 assuming parameters corresponding to a range of observational facilities. As before six cases are shown, corresponding to Lyα surveys with areas of $A_{\rm sky} = 10$ and 100 square degrees performed using a 2m telescope with 1 hour integrations; combined with low-frequency arrays of collecting area corresponding to 1, 10 and 100 LFDs with an integration time 1000 hours. Only one panel is shown because only space based observations have been considered, and because the SN calculated using equation (29) is not dependent on the parameter $f_{\rm flat}$.

Figure 5 shows that at angles below a few arc-minutes, the cross-correlation will be detectable using space-based wide-field imaging (~ 10 square degrees) combined with a low-frequency-array collecting area of at least 10 times that of the LFD. At large scales the signal to noise ratio is substantially improved relative to Figure 4. Figure 5 shows that space-based near-IR surveys with areas of 100 square degrees could achieve $SN \sim 5$ on angles of $5 - 10'$ when combined with the LFD.

In principle, the removal of atmospheric skyglow could also be accomplished through subtraction of independent regions of sky in analogy to equations (27-29). However since (unlike the zodiacal light) the atmospheric skyglow is variable on short timescales, the removal would need to be averaged over a large number of pointings. We have not attempted to compute the signal-to-noise for a detection in this case.

9 DISCUSSION

Standard models for stellar reionization of the IGM predict a clear anti-correlation between the distribution of bright resolved galaxies and the 21cm signal (Wyithe & Loeb 2006). Moreover this anti-correlation would be easily detectable (Wyithe & Loeb 2007; Furlanetto & Lidz 2007). However reionization is thought to be dominated by low mass galaxies. In this paper we have demonstrated that the cross-correlation between fluctuations in the surface brightness of Lyα emission (as a proxy for star formation rate) and the redshifted 21cm signal, will directly test the existence of a causal link between the production of ionizing photons by stars and the reionization of the IGM.

The faint galaxies that make up the unresolved component of high redshift emission produce most of the Lyα emission (and corresponding UV radiation). One might therefore suppose that it should be easier to detect the correlation between the unresolved component and the 21cm emission. However (once a redshift is measured) fluctuations in the resolved galaxy distribution do not suffer from extragalactic foreground or sky brightness contamination, while the unresolved emission must be separated from the fluctuating foreground statistically based on its cross-correlation with the 21cm signal. The correlation of the 21cm signal with a fluctuating Lyα surface brightness will therefore be substantially more difficult to detect than a correlation with resolved galaxies (Wyithe & Loeb 2007; Furlanetto & Lidz 2007).

In this paper we have assumed that measurement of the cross-correlation between 21cm intensity and diffuse Lyα emission would be performed using observations in a narrow near-IR band. The advantage of a narrow band is that the relative fluctuations in both Lyα and 21cm emission are

[4] Note that there is still a contribution to $\sigma_{\rm flat}$ resulting from the extra-galactic foreground. However this contribution is ~ 100 times smaller than the contribution from zodiacal light and we ignore it for this calculation

larger than they would be if averaged over a wider line-of-sight interval, corresponding to a broad band. On the other hand, the signal-to-noise ratio in an individual pointing is increased if a broad band is used, making detection of the smaller fluctuations easier. The key ingredient to measuring the fluctuations due to star formation at high redshift is the ability to remove fluctuations due to the foreground galaxies. As we have shown, these fluctuations are comparable in magnitude to the Lyα signal. Existing measurements of fluctuations in unresolved emission have required subtraction of an estimated fluctuating foreground component (Kashlinsky et al. 2005). Here we suggest removal of the foreground fluctuations statistically using the fact that these are uncorrelated with the redshifted 21cm emission. The 21cm emission also has a fluctuating foreground, and this foreground will be correlated with the foreground in the Lyα observations. It is proposed as part of upcoming 21cm experiments, that this redshifted 21cm foreground be removed using the smoothness of the spectrum of foreground sources (which will be compared with the rapid frequency fluctuations of the 21cm signal). This subtraction method will reveal the narrow-band 21cm fluctuations, but will not allow detection of broad-band fluctuations. Hence a narrow-band near-IR observation would be the optimal choice for detecting the cross-correlation between the Lyα and 21cm signals.

We have analyzed the prospects for detection of the predicted cross-correlation between Lyα and 21cm emission, and found that detection will be possible at angular scales smaller than $\sim 10'$. At scales of $5 - 10'$ the measurement could be performed using near-IR imaging from space, combined with a low frequency array having a collecting area equal to that of the LFD. At smaller angular scales the SN can be significantly increased, but will require a collecting area for redshifted 21cm observations of at least 10 times the LFD. Observations from the ground will be limited by the difficulties of subtracting a sufficiently flat sky, which would be dominated by atmospheric sky-glow. When 21cm observations are combined with space-based near-IR imaging, the detection on scales smaller than a few arc-minutes will be limited by the sensitivity to the 21cm signal. This will be true even when the experiment is performed with a small aperture optical telescope over a moderate field of view (~ 10 square degrees).

A futuristic wide-field space-based survey telescope combined with a Square-Kilometer-Array would detect the cross-correlation at very high signal-to-noise ratios over a range of angular scales below a few arc-minutes. The space-based near-IR survey need not have a very highly sampled point-spread function beyond that necessary for the subtraction of resolved galaxies. A space based survey telescope like the proposed Supernova Acceleration Probe (SNAP) would therefore provide the ideal facility with which to explore the connection between star formation and the reionization of the universe. To perform an experiment of the sort proposed in this paper, the survey telescope would need to carry an appropriate narrow-band filter. To study star formation at $z \sim 7$, corresponding to the examples presented in this paper the filter should be centered at a wavelength of $\sim 9700\AA$ with a width of $\sim 100\AA$.

Acknowledgments The research was supported by the Australian Research Council (JSBW & BPS) and Harvard University grants (AL). JSBW acknowledges the hospitality of the Institute of Astronomy at Cambridge University where this work was completed.

REFERENCES

Barkana, R., & Loeb, A. 2001, Phys. Rep., 349, 125
Barkana, R., & Loeb, A. 2005, ApJL, 624, L65
Bernstein, R. A., Freedman, W. L., & Madore, B. F. 2002, ApJ, 571, 56
Bharadwaj, S., & Ali, S. S. 2005, MNRAS, 356, 1519
Biermann, P. L., & Kusenko, A. 2006, Physical Review Letters, 96, 091301
Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440
Dijkstra, M., Haiman, Z., Rees, M. J., & Weinberg, D. H., Astrophys. J., 601, 666-675 (2004)
Dijkstra, M., Wyithe, J.S.B, Haiman, Z. 2007, MNRAS, sumbitted
Dijkstra, M., Wyithe, J.S.B, 2007, MNRAS, submitted
Dijkstra, M., Lidz, A., Wyithe, J.S.B, 2007, MNRAS, submitted
Efstathiou, G., Mon. Not. R. Astron. Soc., 256, 43-47 (1992)
Fan, X., et al. 2006, AJ, 132, 117
Furlanetto, S., & Lidz, A. 2007, ArXiv Astrophysics e-prints, arXiv:astro-ph/0611274
Hansen, S. H., & Haiman, Z. 2004, ApJ, 600, 26
Kashlinsky, A., Arendt, R. G., Mather, J., & Moseley, S. H. 2005, *Nature*, 438, 45
Kashlinsky, A., Arendt, R. G., Mather, J., & Moseley, S. H. 2007, ApJL, 654, L5
Kasuya, S., & Kawasaki, M. 2007, Journal of Cosmology and Astro-Particle Physics, 2, 10
Leinert, C., et al. 1998, Astron. Astrophys. Supp., 127, 1
Leitherer, C., et al. 1999, ApJS, 123, 3
Loeb, A. 2006, "First Light", lecture notes for the SAAS-Fee Winter School, April 2006 (to be published by Springer Verlag); ArXiv Astrophysics e-prints, arXiv:astro-ph/0603360
Madau, P., Rees, M. J., Volonteri, M., Haardt, F., & Oh, S. P. 2004, ApJ, 604, 484
McQuinn, M., Hernquist, L., Zaldarriaga, M., & Dutta, S. 2007, ArXiv e-prints, 704, arXiv:0704.2239
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347
Morales, M. F., Bowman, J. D., & Hewitt, J. N. 2006, ApJ, 648, 767
Press, W., Schechter, P., 1974, ApJ., 187, 425
Quinn, T., Katz, N., & Efstathiou, G., 278, L49-L54 (1996)
Ricotti, M., Ostriker, J. P., & Gnedin, N. Y. 2005, MNRAS, 357, 207
Ripamonti, E., Mapelli, M., & Ferrara, A. 2007, MNRAS, 375, 1399
Spergel et al., 2006, ArXiv Astrophysics e-prints, arXiv:astro-ph/0603449
Stark, D. P., Loeb, A., & Ellis, R. E. 2007, ArXiv Astrophysics e-prints, arXiv:astro-ph/0701882
Thoul, A. A., & Weinberg, D. H., Astrophys. J., 465, 608-116 (1996)
White, R., Becker, R., Fan, X., Strauss, M., 2003, Astron J., 126, 1
Wyithe, J. S. B., & Loeb, A. 2007, MNRAS, 375, 1034

Wyithe, J. S. B., Loeb, A., Barnes, D.G., 2005, ApJ, 634, 715
Wyithe, J. S. B., Morales, M., 2007, MNRAS, submitted